



Principes de transformations sensorimotrices à l'aide de neurones Sigma-Pi

Jérémy Fix

► To cite this version:

Jérémy Fix. Principes de transformations sensorimotrices à l'aide de neurones Sigma-Pi. NeuroComp : 1ère Conférence Francophone de Neurosciences Computationnelles, Oct 2006, Pont-à-Mousson, France. inria-00166537

HAL Id: inria-00166537

<https://inria.hal.science/inria-00166537>

Submitted on 7 Aug 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PRINCIPES DE TRANSFORMATIONS SENSORIMOTRICES À L'AIDE DE NEURONES SIGMA-PI

Jeremy Fix

Loria, Campus Scientifique, BP 239
54506 Vandoeuvre-les-Nancy, France
email: Jeremy.Fix@Loria.fr

Résumé

Le calcul des diverses transformations sensorimotrices effectuées par le cerveau requiert généralement la combinaison de flux sensoriels de différentes natures (e.g. visuelles, auditives, proprioceptives, etc.) qui permettent au final de construire des représentations pertinentes et adaptées à l'action qui doit être engagée [1]. Ainsi, lorsque l'on souhaite saisir un objet présent dans le champ visuel, les informations visuelles combinées aux informations proprioceptives en provenance de différents effecteurs (coude, épaule, cou, etc.) permettent de déterminer précisément et à chaque instant la position de l'objet par rapport à un référentiel moteur propre à l'effecteur engagé dans l'action, en l'occurrence, le bras dans l'exemple donné.

Considérant un stimulus visuel quelconque (avec l'hypothèse qu'il peut être réduit à un point de l'espace) présent au niveau de la surface rétinienne, il est possible d'attacher un nombre important de représentations dans différents référentiels. Le premier référentiel (et le plus évident en raison de la rétinopie) est celui lié à l'oeil, qui permet de savoir où se trouve le stimulus par rapport au point de fixation actuel. Il s'agit là d'un référentiel sensoriel au sens où la position du stimulus est identifiée avec sa perception sur la surface rétinienne. Il est également possible de considérer la position de ce stimulus dans le référentiel moteur de l'oeil en considérant cette fois le mouvement moteur qu'il est nécessaire d'effectuer pour le centrer sur la rétine, ce mouvement pouvant être approximé par la position sensorielle du stimulus.

Nous souhaitons présenter ici un mécanisme computationnel permettant de combiner une information de type visuel (la position des stimuli visuels dans un référentiel moteur lié à l'oeil) et une information de type proprioceptif (la position de l'oeil dans l'orbite) permettant de construire simplement une représentation dite centrée tête (indépendante de la position des yeux dans leur orbite) qui se trouve être une représentation dans un référentiel moteur lié à la tête. Nous sommes conscients qu'il s'agit là d'un vieux problème et qu'il existe de nombreux travaux et modèles ([2, 3]) permettant d'établir de telles représentations. Ceux-ci reposent néanmoins sur la construction d'une représentation intermédiaire

sous la forme d'une carte associative dont le cardinal se trouve généralement être le produit des cardinaux des cartes d'entrée. En d'autres termes, les auteurs travaillent avec des entrées à une dimension et les dimensions de la carte associative grandissent avec le carré des dimensions des cartes d'entrée. Le passage naïf à des informations en deux dimensions conduit donc rapidement à une explosion combinatoire de la taille des représentations intermédiaires. Nous proposons ici d'éviter cet écueil par l'utilisation de neurones de type sigma-pi originellement introduits par [4] (voir également [5]) de telle sorte que la transformation combinant les informations sensorielles et proprioceptives soit représentée directement via la topologie des connections.

Le modèle que nous proposons se place dans le cadre du paradigme des champs neuronaux discrets à deux dimensions ([6]) qui repose sur une équation différentielle régissant la dynamique des neurones. Les informations visuelles et proprioceptives sont représentées sous forme de paquets d'activités possédant un profil gaussien. La représentation centrée tête est alors obtenue par un calcul distribué du produit de convolution des entrées.

Mots clefs : Transformations sensorimotrices, neurones Sigma-Pi, champs neuronaux.

1 Contexte

Comme nous l'avons indiqué en introduction, nous nous plaçons dans le cadre du paradigme des champs neuronaux discrets, dans lequel les unités de calculs, appelées neurones, caractérisées par leur activité notée u , sont regroupées en carte. Une carte est un ensemble de neurones, auxquels est associé une coordonnée à une ou deux dimensions que nous noterons de manière générale \mathbf{x} . Chaque neurone peut être connecté à des neurones d'autres cartes par des connections afférentes ou à des neurones de la même carte par des connections latérales. Si on note $w(\mathbf{x}, \mathbf{x}')$ le poids d'une connection latérale entre un neurone \mathbf{x}' et un neurone \mathbf{x} et $s(\mathbf{x}, \mathbf{y})$ le poids d'une connection afférente entre un neurone \mathbf{y} et un neurone \mathbf{x} , l'évolution de l'activité du neurone \mathbf{x} est régie par l'équation différentielle du premier ordre (1). En général, les fonctions de poids $s(\mathbf{x}, \mathbf{y})$ et $w(\mathbf{x}, \mathbf{x}')$ sont des gaussiennes ou des différences de gaussiennes.

$$\tau \cdot \frac{\partial u(\mathbf{x}, t)}{\partial t} = -u(\mathbf{x}, t) + \sum_{\mathbf{x}' \in M} w(\mathbf{x}, \mathbf{x}') u(\mathbf{x}', t) + \sum_{\mathbf{y} \in M'} s(\mathbf{x}, \mathbf{y}) \cdot u(\mathbf{y}, t) \quad (1)$$

Cette équation différentielle du premier ordre est une version discrétisée de l'équation d'évolution proposée dans la théorie des champs neuronaux continus (*Continuum Neural Field Theory*) proposée par [7] dans le cas de cartes à une dimension et étudiée par [6] dans le cas à deux dimensions. Dans cette équation apparait clairement le modèle standard de neurone sommateur. Néanmoins, d'autres alternatives pour intégrer les entrées d'un neurone peuvent être considérées. En particulier, [4] introduit des neurones Sigma-Pi dont l'équation d'évolution diffère de la précédente uniquement sur la façon dont le neurone intègre ses entrées :

$$\tau \cdot \frac{\partial u(\mathbf{x}, t)}{\partial t} = -u(\mathbf{x}, t) + \sum_M w_M(\mathbf{x} - \mathbf{x}') u(\mathbf{x}', t) + \sum_{i \in I} w_i \cdot \prod_{\mathbf{y} \in M_i} u(\mathbf{y}, t) \quad (2)$$

2 Les approches utilisant une représentation intermédiaire

Une transformation sensorimotrice est une fonction non linéaire qui associe une représentation motrice à une information sensorielle; Pour saisir avec la main un objet perçu visuellement, il est nécessaire de transformer l'information sensorielle de la projection de l'objet sur la rétine en une information motrice qui permette de guider le mouvement du bras en direction de la cible. Il est montré dans la littérature que certains neurones du cortex pariétal coderaient des informations dans des référentiels associés à

différents effecteurs comme l'oeil, la tête, le bras, etc. ([1]). Le cerveau construirait alors différentes représentations en intégrant successivement les informations proprioceptives associées à la position de l'oeil par rapport à la tête, de la tête par rapport au tronc, etc. (cf figure 1). Dans [8] (voir également [9]), les auteurs proposent une alternative qui consiste à construire une unique représentation intermédiaire, à l'aide de fonctions à base radiale qui combinent toutes les informations proprioceptives et qui est utilisée pour construire les représentations associées aux effecteurs. Néanmoins, comme le souligne les auteurs, une unique représentation intermédiaire nécessite énormément de neurones. En effet, dans le cas simple de la construction d'une représentation centrée sur la tête, à partir des informations visuelles dans une représentation rétinotopique, utilisant également la position de l'oeil dans le globe oculaire, les dimensions de la représentation intermédiaire grandissent avec le carré des dimensions des cartes d'entrée. L'utilisation d'une telle représentation intermédiaire conduit donc à une explosion combinatoire du nombre d'unité utilisées, et ce, alors qu'une seule information proprioceptive est intégrée. Dans [2], les auteurs proposent une version nuancée de leur approche dans laquelle plusieurs représentations intermédiaires, utilisant des fonctions à base radiale, intègrent successivement différentes informations proprioceptives : pour chaque information proprioceptive, une représentation intermédiaire est construite. Les travaux de [10], [11] et [12] utilisent également une représentation intermédiaire pour calculer des transformations sensorimotrices, afin de reproduire les réponses multiplicatives observées pour des neurones du cortex pariétal, en utilisant exclusivement des neurones sommateurs.

3 Exemple d'utilisation des neurones sigma pi

Afin d'éviter l'explosion combinatoire du nombre de neurones utilisés dans une représentation intermédiaire, nous proposons de nous passer de cette représentation en utilisant des neurones sigma-pi et en représentant la transformation uniquement grâce à la topologie des connections. Pour illustrer l'intérêt des neurones sigma-pi dans le calcul de telles transformations, considérons l'exemple de la construction d'une représentation centrée sur la tête à partir d'une représentation rétinotopique et de la position de l'oeil dans le globe oculaire. Nous faisons l'hypothèse que ces trois informations sont représentées dans des cartes de neurones à deux dimensions. Nous faisons également l'hypothèse que la position de l'oeil est représentée par un profil d'activité gaussien et que la représentation centrée sur l'oeil contient un ensemble de bulles d'activité de profil gaussien, centrées sur chaque stimulus visuel. Les neurones sigma-pi de la représentation centrée sur la tête combinent alors ces deux entrées selon l'équation (3), adaptée de l'équation (2). Pour des raisons de convention (pour

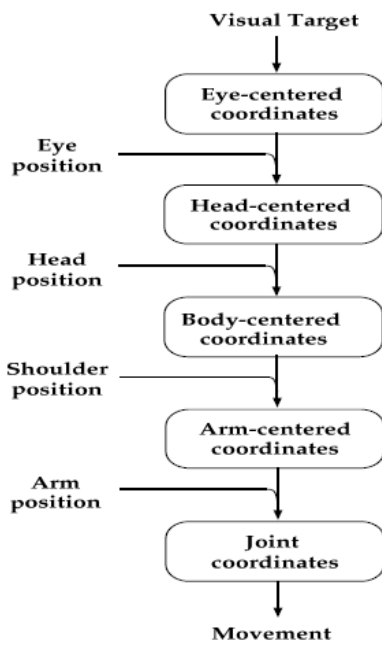


Figure 1. Pour orienter le bras vers une cible visuelle, le cerveau construirait plusieurs représentations en intégrant successivement différentes informations proprioceptives. Adapté de [8]

représenter la position *oeil centrée* par une gaussienne centrée dans la représentation de la position de l'oeil, il est nécessaire de faire apparaître un terme d'offset dans l'écriture du produit de convolution mais nous ne l'avons pas fait apparaître pour ne pas allourdir l'équation, le principe restant le même.

$$\tau \cdot \frac{\partial hc(x, t)}{\partial t} = -hc(x, t) + \sum_{y \in \mathbb{R}^2} ep(y, t) \cdot ec(x - y, t) \quad (3)$$

où $hc(\mathbf{x}, t)$ représente l'activité du neurone \mathbf{x} de la représentation centrée sur la tête (avec $\mathbf{x} \in \mathbb{R}^2$) à l'instant t , $ep(\mathbf{y}, t)$ l'activité du neurone \mathbf{y} de la représentation de la position de l'oeil à l'instant t et $ec(\mathbf{z}, t)$ l'activité du neurone \mathbf{z} de la représentation rétinotopique à l'instant t . La représentation centrée sur la tête est ainsi calculée par le produit de convolution de la représentation centrée sur l'oeil et de la position de l'oeil. D'une part le produit de convolution est une opération linéaire et d'autre part, le produit de convolution de deux gaussiennes est une gaussienne dont le centre peut être obtenu facilement en sommant la position des centres des gaussiennes convoluées. Or, dans une première approximation, la position d'un stimulus dans un référentiel centré sur la tête peut être obtenu en sommant la position de ce stimulus dans le référentiel rétinotopique et la position de l'oeil par rapport à la tête.

La figure 2 illustre le fonctionnement du mécanisme. L'activité des unités de la représentation centrée sur la tête

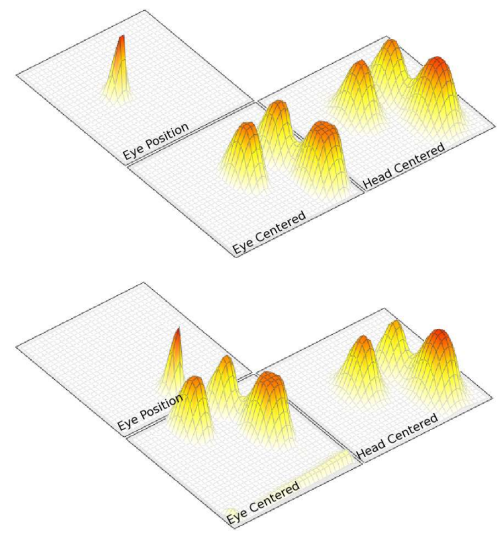


Figure 2. L'entrée visuelle est constitué de trois stimuli représentés par des paquets d'activité de profil gaussien centrés sur leur position. En haut : Pour une position de l'oeil donnée, la représentation centrée sur la tête est obtenue en convoluant la représentation centrée sur l'oeil et la position de l'oeil. En bas : pour une seconde position de l'oeil, la perception visuelle est modifiée mais la représentation centrée sur la tête reste la même.

satisfait l'équation 3. Ainsi la représentation centrée sur la tête est obtenue en convoluant la représentation centrée sur l'oeil et la position de l'oeil. Il est intéressant de constater que lorsque l'oeil bouge (de la figure 2 en haut à la figure 2 en bas), la représentation centrée sur la tête est invariante.

En restant dans l'exemple du passage d'une représentation rétinotopique à une représentation centrée sur la tête, il est intéressant d'étudier la taille de l'espace représenté et le nombre de neurones alloués à cette représentation. L'équation 3 présente le principe du calcul et peut être adapté pour conserver un nombre constant de neurones par représentation alors même que l'espace couvert par un mouvement de la tête est plus important que l'espace couvert par un mouvement de l'oeil. Ceci peut être réalisé simplement en introduisant des coefficients de proportionnalité dans la détermination des couples de neurones dont on fait le produit lors de l'intégration des entrées des neurones sigma-pi. En notant EF le rapport entre le nombre de neurones de la représentation centrée sur l'oeil et l'espace couvert par un mouvement de l'oeil, selon une dimension et HF ce même rapport calculé pour la tête, l'équation 3 peut être adaptée en l'équation 4 afin de tenir compte de l'espace couvert par chaque effecteur.

$$\tau \cdot \frac{\partial hc(x, t)}{\partial t} = -hc(x, t) + \sum_{y \in \mathbb{R}^2} ep(y, t) \cdot ec\left(\frac{EF}{HF} \cdot x - y, t\right) \quad (4)$$

La figure 3 illustre le fonctionnement du mécanisme

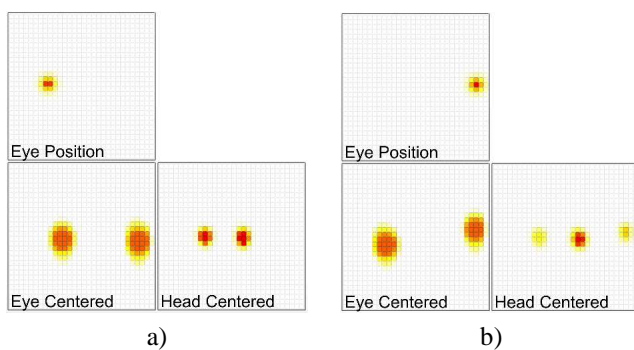


Figure 3. La scène est constituée de trois stimuli alignés horizontalement. a) Un premier mouvement oculaire permet d'en observer deux, qui sont alors représentés dans la représentation centrée sur la tête. b) Un second mouvement oculaire vers la droite permet d'observer le troisième stimulus. La représentation centrée sur la tête couvre un espace plus grand que la représentation centrée sur l'oeil. Les neurones représentant le stimulus le plus à gauche n'étant plus excités par la perception, voient leur activité tendre vers zéro.

avec l'équation précédente. On constate alors que l'espace couvert par la représentation centrée sur la tête est plus grand que celui couvert par la représentation centrée sur l'oeil, alors que le nombre de neurones de chaque représentation est le même; en quelque sorte, on observe un effet de zoom, ou, dit autrement, les neurones de la représentation centrée sur la tête ont des champs récepteurs plus grands que les neurones de la représentation centrée sur l'oeil. Pour la simulation de la figure 3, trois stimuli sont alignés horizontalement. Un premier mouvement oculaire permet d'en observer deux, puis un second mouvement oculaire vers la droite permet d'observer le troisième. Alors qu'à chaque instant seuls deux stimuli sont observables, les trois stimuli peuvent être représentés dans la représentation centrée sur la tête.

4 Discussion

Nous venons de présenter un mécanisme permettant de calculer une transformation sensorimotrice à l'aide de neurones sigma-pi, évitant ainsi l'écueil de l'explosion combinatoire du nombre de neurones utilisés dans un calcul passant par une représentation intermédiaire. Cette solution présente l'avantage d'être moins coûteuse en temps de calcul. Par ailleurs, outre l'extension qui consiste à modifier l'équation 3 pour tenir compte de l'espace à représenter et du support sur lequel la représentation est calculée, une seconde extension envisageable se base sur la remarque suivante : une représentation motrice doit contenir l'ensemble des stimuli atteignables par l'effecteur associé à la représentation. Or, la représentation rétinotopique telle que nous l'avons présentée se limite au champ visuel. Nous pourrions imaginer augmenter les dimensions

de l'espace représenté pour qu'il englobe la globalité des stimuli accessibles par un mouvement de l'oeil. La partie centrale de la représentation serait alimentée par la perception et la représentation centrée sur la tête serait construite comme précédemment. Pour le moment, le mécanisme proposé est purement feedforward. Il est alors intéressant d'introduire des connections feedbacks de la représentation centrée sur la tête vers la représentation centrée sur l'oeil, pour compléter la représentation rétinotopique afin qu'elle ne se limite plus simplement au champ visuel, ce calcul se faisant également à l'aide de neurones sigma-pi.

References

- [1] C.L.Colby. Action-oriented spatial reference frames in cortex. *Neuron*, 20:15–24, 1998.
- [2] A. Pouget, S. Deneve, and J.R. Duhamel. A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews*, 3(9):741–748, 2002.
- [3] P. Baraduc, E. Guigon, and Y. Burnod. Where does the population vector of motor cortical cells point during reaching movements ? *NIPS*, pages 83–89, 1998.
- [4] D.E. Rumelhart, G.E. Hinton, and J.L. McClelland. A general framework for parallel distributed processing. In The MIT Press, editor, *Parallel Distributed Processing, Vol. 1*, pages 45–76, 1986.
- [5] R.J. Williams. The logic of activation functions. In The MIT Press, editor, *Parallel Distributed Processing, Vol. 1*, 1986.
- [6] J.G. Taylor. Neural bubble dynamics in two dimensions. *Biological Cybernetics*, 80:5167–5174, 1999.
- [7] S.I. Amari. Dynamical study of formation of cortical maps. *Biological Cybernetics*, 27:77–87, 1977.
- [8] A. Pouget and T.J. Sejnowski. Spatial transformations in the parietal cortex using basis functions. *Journal of Cognitive Neuroscience*, 9(2):222–237, 1997.
- [9] A. Pouget and L.H. Snyder. Computational approaches to sensorimotor transformations. *Nature neuroscience*, 3:1192–1198, november 2000.
- [10] E.L. Sauser and A.G. Billard. Three-dimensional frames of references transformations using recurrent populations of neurons. *Neurocomputing*, 64:5–24, 2005.
- [11] P. Baraduc and E. Guigon. Population computation of vectorial transformations. *Neural Computation*, 14(4):845–871, 2002.
- [12] E. Salinas and L.F. Abbott. Transfer of coded information from sensory to motor networks. *Journal of Neuroscience*, 15:6461–6474, 1995.